

Linked in.

Career Explorer Data & Architecture

http://career.linkedin.com

Jeffrey Schang Senior Software Engineer

Stanford Chapter of ACM November 12, 2010



- Features
- Architecture
- Data
- Challenges

Jeffrey Schang

Senior Software Engineer, Search and Analytics, LinkedIn

San Francisco Bay Area | Internet

Public Profile: http://www.linkedin.com/in/jeffreyschang

Connections: 159 Email: jschang@linkedin.com







- Students/career changers want to know what they can do
- They want to make informed decisions





- Needs of student
- What if?
 - You knew what type of job people like you get
 - And where they go afterwards

LinkedIn has the data!



- Path Chooser
- Statistical Trends
- Job Listings
- Group Discussions
- Who can help you get where you want to be?





- Path Chooser
 - Major
 - Degree
 - Industry

- University
- Location
- Years of experience





Linked in Go to LinkedIn Recruiter »		* Jeffrey Schang Add Connections
Home Profile Contacts Groups Jobs Inbox 3	Companies More People *	Advanced
Career Explorer beta		
Career Path-1 × Career Path-2 Career Path	-3 Make another path	
A Your career path starting point ×		
What is your major?		
Computer Engineering		
Expected degree?	Males and this is a	· · · · · · · · · · · · · · · · · · ·
Master's \$	Make sure this info is correct + click	
Expected industry?	'Start my path'	
(Internet \$		
Start your path		
You are here!		
	2	
		3





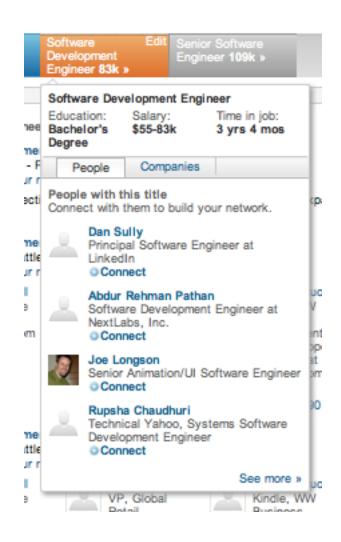


Linked [Go to LinkedIn Recruiter »		* Jeffrey Schang Add Connections
Home	Profile Contacts Groups Jobs Inbox	8 Companies More People -	Q Advanced
	Explorer beta er Path-1 Career Path-2 * Career Pat	Software Engineer Time in job Education Salary 3 yrs 5 mos Bachelor's \$54-82k	
P	Student » Software Computer Development Science, Engineer 83k »	Senior Software Engineer Time in job Education Salary 3 yrs 10 mos Bachelor's \$79-109k	
	Now choose your next	Project Manager Time in job Education Salary 4 yrs Bachelor's \$55-87k	
	step. These are common titles for this stage in your career.	Web Developer Time in job Education Salary 2 yrs 3 mos Bachelor's \$41-63k	?
		Q Search for a title I'm done setting up this path Launch the dashboard now »	5

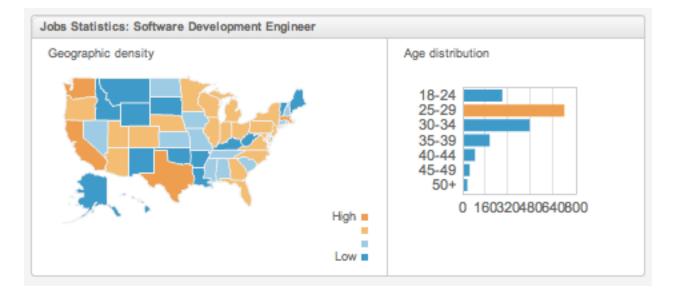


Linked 🛅 Go to LinkedIn Recruiter »				* Jeffrey	y Schang Add	Connections
Home Profile Contacts Groups Jo	bs Inbox 38	Companies More	People *		٩	Advanced
Career Explorer beta Career Path-1 Career Path-2 *	Career Path-3	Make another p	Software Eng Time in job 3 yrs 5 mos	Education	Salary \$54-82k	
Student » Computer Science, Enginee	e g ment E	Senior Software Engineer 109k »	Project Mana Time in job 4 yrs	-	Salary \$55-87k	$\mathbf{\Sigma}$
	Choose one		Software Dev Time in job 3 yrs 1 mo	Education Bachelor's	Salary \$54-82k	
	- we're almost done!		Principal Sof Time in job 5 yrs 3 mos	Itware Engine Education Bachelor's	er	
\bigcirc		Q Search for a title				
				ing up this pa ashboard now		











Software Development Engineer Reading List

ReadingList^w amazon



Rework See this book on Amazon »

See this book on Amazon »

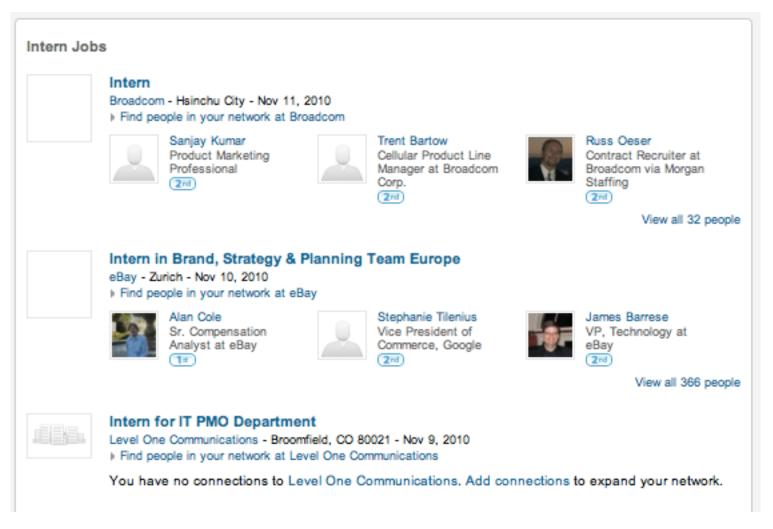




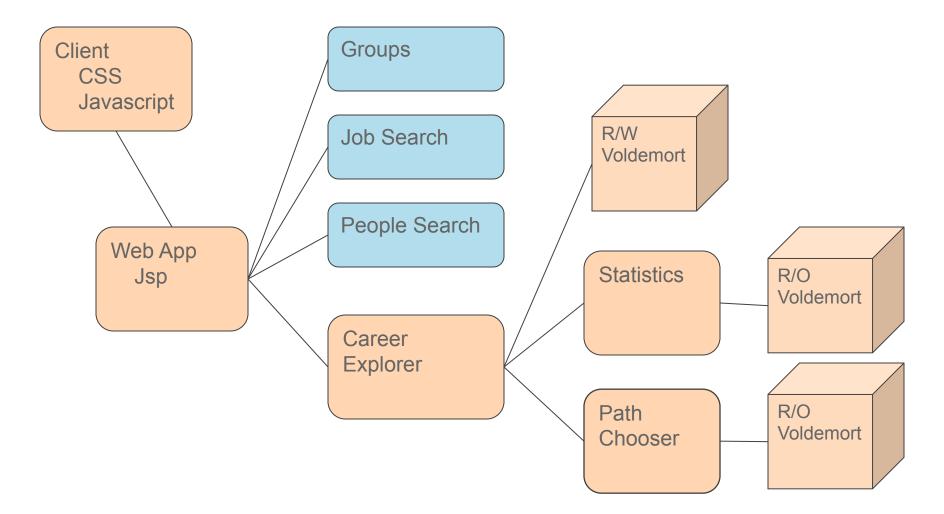
Good to Great: Why Some Companies Make the Leap... and Others Don't See this book on Amazon »

Outliers: The Story of Success











- No two people are alike
 - Different majors
 - Different industries
 - Different experiences

How to forecast?

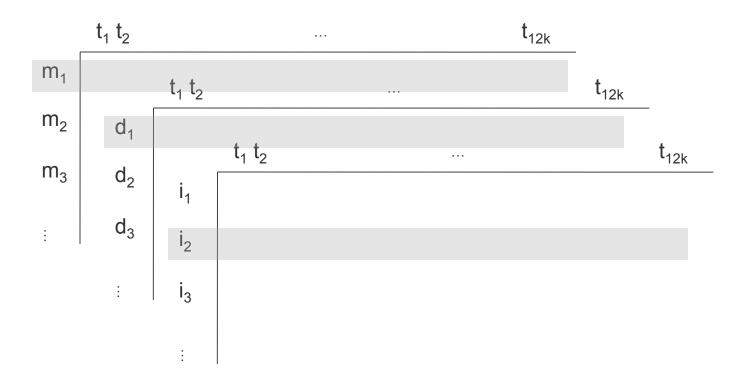




- We choose a naïve Bayesian statistical model combine probabilities of up to 7 parameters:
 - 1) For a given parameter, calculate the probability of the next job (about 12,000 standardized job titles)
 - 2) For each job title, multiply the probabilities of all the parameters and normalize

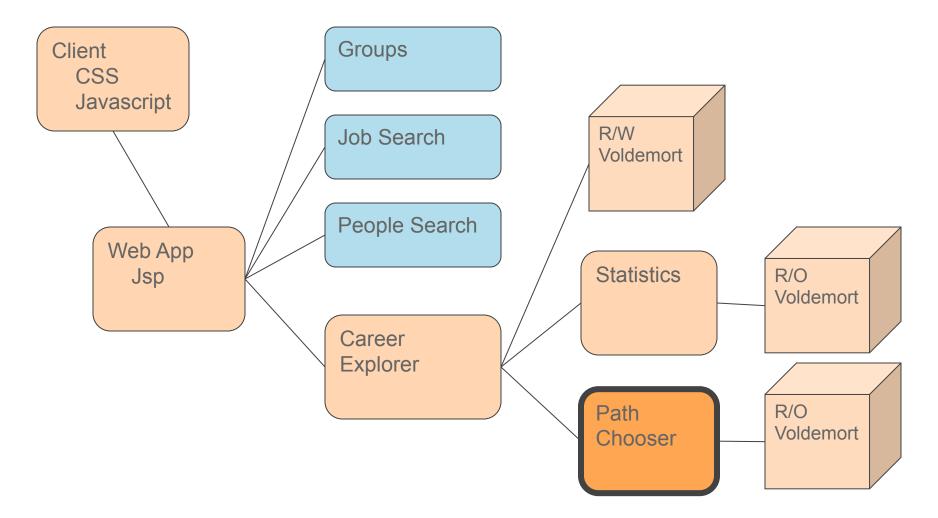






Processing: multiply up to 7 rows of 12k items, < 20 ms





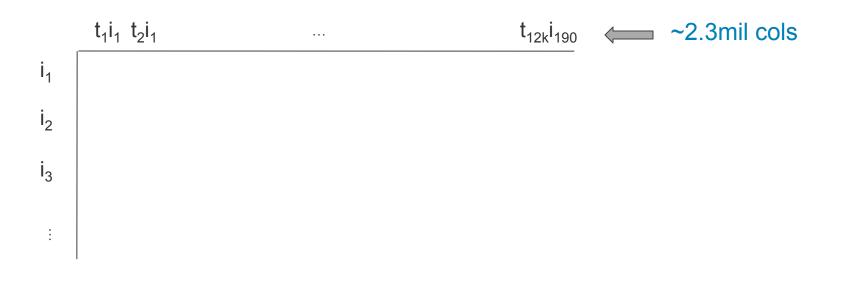


- Are the forecasts realistic?
 - Often top titles are generic title (e.g. Intern, Account Manager)
 - Then next title is also generic (e.g. Owner, V.P.)
 - Overweighting of technology professionals
 - Generic title is specific to industry





- How to make forecasts realistic?
 - Title is actually a compound key with industry





- 2.3 million columns per parameter:
 - Now too much computation per request
 - I/O per request
 - Also size of data tables getting expensive





- Only use the compound key for generic titles that span many industries
 - Reduces columns to approximately 240k
 - Processing per request: 15 ms I/O, 35 ms calculation





- Groundbreaking tool for career planning unlike anything else out there
- Made possible through the combined data of LinkedIn's 80,000,000 members





Linked in.

Career Explorer Data & Architecture

http://career.linkedin.com

Jeffrey Schang

Senior Software Engineer, Search and Analytics, LinkedIn

San Francisco Bay Area | Internet

Public Profile: http://www.linkedin.com/in/jeffreyschang

Connections: 159 Email: jschang@linkedin.com

